

# On the relation between kinetically constrained models of glass dynamics and the random first-order transition theory

Laura Foini,<sup>1</sup> Florent Krzakala,<sup>2</sup> and Francesco Zamponi<sup>3</sup>

<sup>1</sup>*Université Pierre et Marie Curie – Paris VI, LPTHE,*

*Tour 13 5ème étage, 4 Place Jussieu, 75252 Paris Cedex 05, France*

<sup>2</sup>*ESPCI ParisTech, CNRS UMR 7083 Gulliver, 10 rue Vauquelin, 75005 Paris, France*

<sup>3</sup>*LPT, École Normale Supérieure, CNRS UMR 8549, 24 Rue Lhomond, 75005 France*

In this paper we revisit and extend the mapping between two apparently different classes of models. The first class contains the prototypical models described –at the mean-field level– by the Random First Order Transition (RFOT) theory of the glass transition, called either “random XORSAT problem” (in the information theory community) or “diluted  $p$ -spin model” (in the spin glass community), undergoing a single-spin flip Glauber dynamics. The models in the second class are Kinetically Constrained Models (KCM): their Hamiltonian is that of independent spins in a constant magnetic field, hence their thermodynamics is completely trivial, but the dynamics is such that only groups of spin can flip together, thus implementing a kinetic constraint that induces a non-trivial dynamical behavior. A mapping between some representatives of these two classes has been known for long. Here we formally prove this mapping at the level of the master equation, and we apply it to the particular case of Bethe lattice models. This allows us to show that a RFOT model can be mapped exactly into a KCM. However, the natural order parameter for the RFOT model, namely the spin overlap, is mapped into a very complicated non-local function in the KCM. Therefore, if one were to study the KCM without knowing of the mapping onto the RFOT model, one would guess that its physics is quite different from the RFOT one. Our results instead suggest that these two apparently different descriptions of the glass transition are, at least in some case, closely related.

## I. INTRODUCTION

The Random First-Order Transition (RFOT) [1–4] and the dynamic facilitation theory based on Kinetically Constrained Models (KCM) [5–7] are often viewed as alternative pictures of glass formation [6, 8–10]. Indeed, in the former the glass transition is induced by an underlying thermodynamic transition, the so-called “Kauzmann transition”, while in the latter the thermodynamics is assumed to be completely trivial, and the glass transition is a purely dynamical arrest induced by complex kinetic constraints<sup>1</sup>.

The RFOT theory<sup>2</sup> is based on the analysis of a particular class of mean-field spin glass models with  $p$ -spin interactions [1]. It has been known for long [5, 11, 12] that some particular finite-dimensional version of these  $p$ -spin models (the so-called “plaquette models”) can be explicitly mapped onto KCMs. However, the physics of these models, at least in two dimensions, seems to be

quite different from the RFOT one [5, 11–13], thereby preventing one to establish a direct mapping between the two approaches.

The aim of this work is to discuss an example in which, on the contrary, a KCM can be exactly mapped into a model described by the RFOT theory. In order to do this, we first generalize the mapping discussed in [5, 11–13] to an arbitrary geometry, with the only requirement that *i*) each spin is involved in exactly  $p$  interactions, *ii*) each interaction involves exactly  $p$  spins, and *iii*) the ground state is unique. By using the method developed by Gosset [14] for quantum Hamiltonians, we show that the Fokker-Plank operator of the  $p$ -spin model can be mapped into the one of the KCM.

We then focus on a Bethe lattice geometry, because this case allows for a very complete and detailed analysis, and in particular the Bethe lattice  $p$ -spin model is exactly described by the RFOT theory. More precisely, we consider the two following models:

- The first model is a spin model defined on a random regular factor graph (or Bethe lattice) geometry, where each spin is randomly connected to three interactions and each interaction involves three spins. The dynamics is a single spin-flip Glauber dynamics. The model is known as “random 3-XORSAT” in the information theory community, and as the diluted 3-spin in the spin glass community. It was studied intensively by both communities [15–19], and it was shown that the dynamical spin-correlations show a plateau in time that is naturally interpreted from a thermodynamical point of view in terms of the emergence of disconnected er-

<sup>1</sup> Note that sometimes (especially in the mathematical literature) a more stringent definition of KCM is employed. Namely, one requires that the dynamics of a finite size system has absorbing (or “blocked”) states. In the following, we will just define a KCM as a system that has a trivial thermodynamics, and a complex dynamic behavior induced by specific dynamic rules.

<sup>2</sup> In this paper, by RFOT theory we denote the ensemble of theories that can be related to the exact solution of mean-field spin glass models (therefore including mode-coupling theory, molecular liquid replica theory, and finite-dimensional extensions). In particular, we will mostly be concerned with mode-coupling-like dynamics, but we will keep calling it RFOT dynamics.

godic states that trap the system's dynamics (see Fig. 1). Rigorous results have in fact shown that for such a system the Gibbs-Boltzmann measure can be decomposed into disconnected components and purely dynamical quantities are recovered from a thermodynamical study [17, 18, 20]. We will refer to this model simply as “XORSAT” in the following.

- The second model is defined on the same geometry as the first one. Its Hamiltonian is given by non-interacting spins in a magnetic field, and its dynamics only allows a simultaneous flipping of a group of three spins connected to the same “interaction”. This model displays the typical behavior of a Bethe lattice KCM [21], i.e. it has trivial equilibrium thermodynamical properties but the dynamic rules lead to a complex dynamical behavior characterized by a finite-temperature dynamical transition. As in other KCMs, despite the simplicity of the equilibrium Gibbs-Boltzmann measure, the dynamical correlations display slow relaxation and characteristics that resemble those of glassy systems. In the following we will refer to this second model as the KCM.

Thanks to the general mapping discussed above, we show that these two models are in fact completely equivalent: the master equations describing their dynamics can be mapped one into the other by a suitable transformation, that maps the groups of three interacting spins (“plaquettes”) of XORSAT into the spins of the KCM. We show that while different observables share a similar dynamical behavior, and in particular a plateau in their time correlations associated with the dynamical transition, they crucially differ once that they are chosen as order parameters in a static computation.

Therefore, some observables fail to identify the ergodicity breaking. Remarkably, this is the case for all local observables in the KCM: detecting ergodicity breaking in the KCM therefore requires the use of non-local (hence very unnatural and complicated) observables. To illustrate this point, we will consider two distinct order parameters for the XORSAT model: the spin overlap  $q_s$  and the plaquette overlap  $q_p$ . We will show that  $q_s$  is a good order parameter, while  $q_p$  is not. The plaquette overlap for XORSAT becomes the spin overlap for the KCM, therefore the spin overlap is not a good order parameter for the KCM. Moreover, the spin overlap for XORSAT becomes a very complicated non-local function for the KCM; therefore, detecting the glass transition by means of a static computation in the KCM requires the use of a very unintuitive order parameter. If one were to study the KCM without knowing of the underlying mapping on XORSAT, it would be impossible to identify the correct order parameter, and one would conclude that the dynamic transition cannot be detected via a static computation, and that the dynamical facilitation picture is completely different from the RFOT one.

Our results suggest that this conclusion might be wrong: at least for the model we investigated, the two scenarios are dual to each other. The same could be true in more general models [22].

The paper is organized as follows. In section II we establish in general the mapping between the Fokker-Planck equations describing the two classes of models. In section III we focus on the Bethe lattice models and present the results for both dynamic correlation functions and the thermodynamic RFOT potential. Finally, in section IV we discuss our results and their implications.

## II. MAPPING $p$ -SPIN MODELS INTO KCMS

In this section we show how to characterize the equilibrium dynamics of the models and how to construct the mapping between the two.

Let us first outline the general Fokker-Planck equation for the stochastic dynamics of  $N$  Ising spins,  $\sigma_i = \pm 1$  and  $i = 1, \dots, N$ , described by a generic energy function  $E(\underline{\sigma})$ . We define  $P(\underline{\sigma}, t)$  the probability that at time  $t$  the system is in the configuration  $\underline{\sigma}$ . Suppose that the elementary dynamics is induced by the flipping of a set of variables  $k = \{i_k^1, i_k^2, \dots, i_k^{|k|}\}$  of size  $|k|$ , with rate  $w(\underline{\sigma}^{(k)}, \underline{\sigma})$ , where  $\underline{\sigma}^{(k)}$  represents the configuration that is obtained from  $\underline{\sigma}$  by the flip of the set  $k$ . Then, the probability  $P(\underline{\sigma}, t)$  satisfies the equation:

$$\frac{\partial P(\underline{\sigma}, t)}{\partial t} = - \sum_k w(\underline{\sigma}^{(k)}, \underline{\sigma}) P(\underline{\sigma}, t) + \sum_k w(\underline{\sigma}, \underline{\sigma}^{(k)}) P(\underline{\sigma}^{(k)}, t), \quad (1)$$

where the sum over  $k$  stems for all possible sets of variables. In order to reach equilibrium the rates have to satisfy the detailed balance condition:

$$w(\underline{\sigma}^{(k)}, \underline{\sigma}) e^{-\beta E(\underline{\sigma})} = w(\underline{\sigma}, \underline{\sigma}^{(k)}) e^{-\beta E(\underline{\sigma}^{(k)})}, \quad (2)$$

and in the following we will adopt the common choice

$$w(\underline{\sigma}^{(k)}, \underline{\sigma}) = w_0 e^{-\frac{\beta}{2}(E(\underline{\sigma}^{(k)}) - E(\underline{\sigma}))}, \quad (3)$$

with  $w_0 = 1$  in order to set the unit of time.

Defining the matrix  $H$  such that  $\langle \underline{\sigma} | H | \underline{\sigma}' \rangle = \sum_k w(\underline{\sigma}^{(k)}, \underline{\sigma}) \delta(\underline{\sigma}, \underline{\sigma}') - \sum_k w(\underline{\sigma}, \underline{\sigma}^{(k)}) \delta(\underline{\sigma}^{(k)}, \underline{\sigma}')$ , and  $|P(t)\rangle$  the probability vector such that  $\langle \underline{\sigma} | P(t) \rangle = P(\underline{\sigma}, t)$ , Eq. (1) can be written as a Schrödinger equation in imaginary time:

$$\frac{\partial |P(t)\rangle}{\partial t} = -H |P(t)\rangle. \quad (4)$$

It is often convenient to introduce the transformation induced by the diagonal matrix  $U$ , such that  $\langle \underline{\sigma} | U | \underline{\sigma}' \rangle = e^{\frac{\beta}{2} E(\underline{\sigma})} \delta(\underline{\sigma}, \underline{\sigma}')$ , which turns Eq. (4) into:

$$\frac{\partial |\tilde{P}(t)\rangle}{\partial t} = -\hat{H} |\tilde{P}(t)\rangle, \quad (5)$$

with  $|\tilde{P}(t)\rangle = U|P(t)\rangle$  and  $\hat{H} = UHU^{-1}$ . The advantage of this formulation is that the matrix  $\hat{H}$  is Hermitian.

Under the condition that the rates satisfy Eq. (2), one can verify that the equilibrium Gibbs-Boltzmann measure  $P(\underline{\sigma}, t) = P_{GB}(\underline{\sigma}) \propto \exp(-\beta E(\underline{\sigma}))$  is a stationary solution of Eq. (1) and it represents the ground state of the Fokker-Planck operator  $H$  in Eq. (4), and the same holds for  $\tilde{P}_{GB}(\underline{\sigma}) \propto \exp(-\beta E(\underline{\sigma})/2)$  and  $\hat{H}$ .

### A. XORSAT model with spin-flip dynamics

We consider a model of  $N$  Ising spins characterized by 3-body interactions with Hamiltonian:

$$E_{XOR}(\underline{\sigma}) = - \sum_{a=1}^M \sigma_{i_a^1} \sigma_{i_a^2} \sigma_{i_a^3} \quad (6)$$

where  $a$  labels the interactions, each one involving a triplet of spins. Such Hamiltonian represents a particular case ( $p = 3$ ) of a general class of spin models with  $p$ -body interactions that in computer science is also known as XORSAT problem.

The graphical representation of the problem can be obtained through the introduction of a “factor graph”, by assigning a “variable node” to each spin  $i$  and an “interaction node” to each interaction  $a$  and connecting every spin to the interaction nodes in which it takes part. Here we assume that each spin is connected to exactly three interactions, hence the connectivity of spin variables is  $c = 3$ . For the Hamiltonian in Eq. (6), the connectivity of each interaction node is  $p = 3$  and the number  $M$  of these nodes must satisfy  $Mp = Nc$ , then  $M = N$ . In particular, from Eq. (6), the ferromagnetic configuration  $\sigma_i = 1 \forall i = 1, \dots, N$  minimizes all the interactions  $a = 1, \dots, M$  and in what follows we will restrict to the case where this is the only ground state (we will come back to this point in section III).

For this model we study the dynamics induced by single spin flips, so the sum over  $k$  in Eq. (1) is simply the sum over all the spins and the rates  $w(\underline{\sigma}^{(i)}, \underline{\sigma})$  connect configurations differing only for a single spin  $\sigma_i$ . Then, according to Eqs. (3) and (6),  $w_{XOR}(\underline{\sigma}^{(i)}, \underline{\sigma}) = e^{-\beta \sum_{a \in \partial i} \sigma_{i_a^1} \sigma_{i_a^2} \sigma_{i_a^3}}$ , where here and in the following  $a \in \partial i$  indicates the neighborhood of  $i$ , i.e. the interactions  $a$  to which the site  $i$  is connected. Performing the operations outlined in the previous section, one arrives to the following form for the Hermitian Fokker-Planck operator that appears in Eq. (5):

$$\hat{H}_{XOR} = \sum_{i=1}^N e^{-\beta \sum_{a \in \partial i} \hat{\sigma}_{i_a^1}^z \hat{\sigma}_{i_a^2}^z \hat{\sigma}_{i_a^3}^z} - \sum_{i=1}^N \hat{\sigma}_i^x, \quad (7)$$

where  $\hat{\sigma}_i^{x,y,z}$  are the standard Pauli matrices.

### B. Kinetically constrained model

We now introduce the kinetically constrained model (KCM), using different notations that will be suitable for the mapping to be discussed in the next section. We consider a system of  $N$  Ising spins  $s_a = \pm 1$  whose energy function reads:

$$E_{KCM}(\underline{s}) = - \sum_{a=1}^N s_a. \quad (8)$$

Due to the non-interacting nature of  $E_{KCM}(\underline{s})$  the thermodynamical properties of the system are independent of the geometry where the system is defined.

In order to specify the constrained dynamics, we assume that the model is defined on a regular random factor graph, with  $p = c = 3$ , exactly as before, and we define the dynamics by the flip of triplets of spins  $\{a_i^1, a_i^2, a_i^3\}$  surrounding the same “interaction” node  $i$  (hence we use a dual notation with respect to the XORSAT model). The set  $k$  in the Fokker-Planck Eq. (1) is now the set of the three spin identified by  $i$  and, following our choice in Eq. (3), the rates read  $w_{KCM}(\underline{s}^{(i)}, \underline{s}) = \exp(-\beta \sum_{a \in \partial i} s_a)$ . The Hermitian evolution operator that enters in the Fokker-Planck Eq. (5) for the KCM can be expressed in terms of Pauli matrices as follows:

$$\hat{H}_{KCM} = \sum_{i=1}^N e^{-\beta \sum_{a \in \partial i} \hat{s}_a^x} - \sum_{i=1}^N \hat{s}_{a_i^1}^z \hat{s}_{a_i^2}^z \hat{s}_{a_i^3}^z. \quad (9)$$

Note that the choice of the indices  $z$  and  $x$  in Eq. (9) is just an arbitrary choice of basis for the representation of the spin configurations and one can safely exchange the two indices.

### C. Mapping

Following the procedure outlined in [14] in the context of quantum spin systems, in this section we describe the mapping between  $\hat{H}_{XOR}$ , defined on a regular graph with  $c = p$  and such that  $E_{XOR}$  has a unique ground state, into  $\hat{H}_{KCM}$  defined on the dual graph of the interaction nodes.

Consider the operator  $\hat{H}_{XOR}$  defined in Eq. (7) for a given graph with a unique (ferromagnetic) ground state. Thanks to this property, for each interaction  $a$ , it exists one and only one configuration of spins  $\underline{\sigma}^a$  minimizing all the interactions but  $a$ . To show this, let us introduce the  $N \times N$  adjacency matrix  $A$  such that  $A_{ai} = 1$  if the spin  $\sigma_i$  enters in the interaction  $a$  and  $A_{ai} = 0$  otherwise, and define  $\underline{y}^a$  such that  $y_i^a = \frac{\sigma_i^a - 1}{2}$ . The matrix  $A$  is invertible because by assumption the system  $A\underline{y} = \underline{0}$  (that is equivalent to the condition of minimizing  $\bar{E}_{XOR}(\underline{\sigma})$ ) has a unique solution, corresponding to  $\underline{y} = \underline{0}$ . Defining the vector  $\underline{e}^a$  such that  $e_b^a = \delta_{a,b}$ , one gets  $\underline{y}^a = A^{-1} \underline{e}^a$ .

Then, one can introduce the following operators, defined on the interaction nodes of the underlying graph of Eq. (7):

$$\begin{aligned}\hat{s}_a^x &= \hat{\sigma}_{i_1 a}^z \hat{\sigma}_{i_2 a}^z \hat{\sigma}_{i_3 a}^z, \\ \hat{s}_a^z &= \prod_{i=1}^N (\hat{\sigma}_i^x)^{y_i^a}.\end{aligned}\quad (10)$$

These are well defined spins operators because they satisfy the correct commutation relations,  $[\hat{s}_a^x, \hat{s}_b^z] = 0$  if  $a \neq b$ ,  $\{\hat{s}_a^x, \hat{s}_a^z\} = 0$  and  $(\hat{s}_a^x)^2 = (\hat{s}_a^z)^2 = \mathbb{I}$ , where  $\mathbb{I}$  is the  $2 \times 2$  identity matrix. Moreover, it turns out that defining  $\underline{e}^i$  such that  $e_j^i = \delta_{i,j}$ , and using the relation

$$A \underline{e}^i = \sum_{a \in \partial i} \underline{e}^a \Rightarrow \underline{e}^i = A^{-1} \sum_{a \in \partial i} \underline{e}^a = \sum_{a \in \partial i} \underline{y}^a \quad (11)$$

the mapping above can be inverted as follows:

$$\begin{aligned}\hat{\sigma}_i^x &= \hat{s}_{a_1 i}^z \hat{s}_{a_2 i}^z \hat{s}_{a_3 i}^z, \\ \hat{\sigma}_i^z &= \prod_{a=1}^N (\hat{s}_a^x)^{y_i^a}.\end{aligned}\quad (12)$$

Then, enforcing the mapping in Eq. (10) and (12) in  $\hat{H}_{XOR}$ , one obtains the operator  $\hat{H}_{KCM}$  of Eq. (9) defined on the dual lattice. Note that here we discussed the mapping for  $p = c = 3$ , but the extension to any other value of  $p = c$  is straightforward [14].

We have therefore shown that  $\hat{H}_{XOR}$  defined on a given graph can be written as  $\hat{H}_{KCM}$  defined on the dual graph by means of the transformation (12), and viceversa. Note that this implies in particular that the ground state of  $\hat{H}_{XOR}$ ,

$$\langle \underline{\sigma} | GS_{XOR} \rangle = \frac{1}{\sqrt{Z}} e^{-\frac{1}{2} \beta E_{XOR}(\underline{\sigma})}, \quad (13)$$

whose square is the Gibbs distribution, can be written as the one of  $\hat{H}_{KCM}$  on the dual graph:

$$\langle \underline{s} | GS_{KCM} \rangle = \frac{1}{\sqrt{Z}} e^{-\frac{1}{2} \beta E_{KCM}(\underline{s})} \quad (14)$$

and moreover the partition functions of the two models (that correspond to the normalization of the ground states) are identical:

$$Z = \sum_{\underline{\sigma}} e^{-\beta E_{XOR}(\underline{\sigma})} = \sum_{\underline{s}} e^{-\beta E_{KCM}(\underline{s})}. \quad (15)$$

### III. BETHE LATTICE MODELS

Here we specialize our discussion to Bethe lattice models. Different definitions of Bethe lattice have been used in the literature, and here we consider as a model of Bethe lattice a *random regular factor graph*: we construct an ensemble of graphs by giving uniform probability to all the

graphs characterized by  $N$  spins and  $M$  interactions such that each spin appears exactly in  $c = 3$  interactions and each interaction connects precisely  $p = 3$  spins, and such that Eq. (6) has a unique ground state (the last condition is needed in order to use the mapping discussed in section II C). We will be interested in the properties of the model in the thermodynamic limit  $N = M \rightarrow \infty$ . In this limit, such graphs have the property of being locally tree-like, as a Bethe lattice should: indeed, typical loops have length of order  $\log N$ . The lack of a finite dimensional geometry make statistical models defined on such lattices intrinsically mean-field like. Most importantly this property allows for an analytical treatment via the so-called “cavity method”, and many results have been derived in this way [17]. The thermodynamical properties of Eq. (6) are self-averaging with respect to the ensemble of the graphs so defined and it turns out that in the case considered here,  $k = c$ , the ground state is non-degenerate with finite probability [23]. Hence, the restriction to graphs that have a unique ground state, that has to be made here in order to derive the mapping, does not alter the main properties of the model. Note that the ensemble of regular random graphs with  $c = p$  is self-dual, hence the KCM will be defined on the same ensemble of graphs. Note also that the two dimensional triangular lattice considered in [5] represents a (very atypical) instance of such ensemble of graphs.

#### A. XORSAT model

The XORSAT model on the random regular factor graph with  $p = c = 3$  follows exactly the RFOT scenario of the glass transition [17, 18, 24]. It is characterized by a dynamic transition at a temperature  $T_d = 0.5098(2)$ , below which ergodicity is broken and the equilibrium dynamical correlation functions of local observables never relax. In the RFOT scenario there is a second temperature  $T_k$  at which a thermodynamic transition to a replica-symmetry broken state takes place [1]. However, here  $T_k = 0$ , because of the mapping onto independent spins that ensures that the free energy is analytic at all temperatures, see Eq. (15); this can also be checked by an explicit replica or cavity computation [15–17, 24, 25].

##### 1. Spin correlations

In the formalism used above the local spin-spin correlation function for the XORSAT problem reads:

$$C_s^{XOR}(t) = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i(t) \sigma_i(0) \rangle, \quad (16)$$

$$\langle \sigma_i(t) \sigma_i(0) \rangle = \langle GS_{XOR} | \hat{\sigma}_i^z e^{-\hat{H}_{XOR} t} \hat{\sigma}_i^z | GS_{XOR} \rangle,$$

where  $\langle \underline{\sigma} | GS_{XOR} \rangle$  is the ground state of  $\hat{H}_{XOR}$  defined in Eq. (13). These correlation functions are shown in

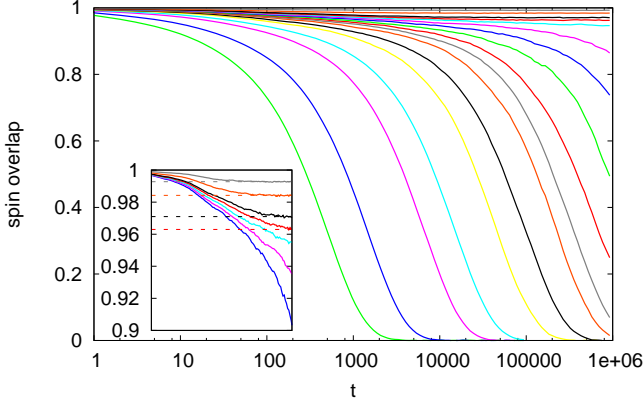


FIG. 1: Local spin-spin correlation function  $C_s^{XOR}(t)$  for the XORSAT model at different temperatures, from left to right  $T=0.7, 0.65, 0.6, 0.58, 0.56, 0.55, 0.54, 0.535, 0.53, 0.525, 0.52, 0.515, 0.51, 0.505, 0.5, 0.48, 0.45$ . The inset shows the comparison of the long time dynamics with the spin overlap that is obtained through the cavity method, which is reported as a dashed line for each temperature [17, 25].

Fig. 1: for  $T < T_d$ , they tend to a plateau  $\bar{q}_s^{XOR}$  at long times. The dynamical correlation functions, here and in the following, are obtained using the Metropolis algorithm for a system of  $N = 60000$  spins. Sampling the initial configuration below  $T_d$  is extremely hard with standard techniques, but it can be easily achieved using the planting technique [20, 26, 27].

The plateau value  $\bar{q}_s^{XOR}$  can be recovered from a statistical treatment of the thermodynamical properties of Eq. (6) via the cavity method [17, 25]. In this treatment, it is assumed that the system at long times decorrelates inside one of the typical ergodic components  $\alpha$  at temperature  $T$ ; then

$$\bar{q}_s^{XOR} = \frac{1}{N} \sum_{i=1}^N \overline{\langle \sigma_i \rangle_\alpha \langle \sigma_i \rangle_\alpha} \quad (17)$$

where the overline indicates the equilibrium average over ergodic components. The equivalence between these two quantities is shown in the inset of Fig. 1.

A more detailed correspondence between statics and dynamics can be established by means of the Franz-Parisi thermodynamical potential [28, 29], defined as:

$$V_s^{XOR}(q) = -\frac{T}{Z} \sum_{\underline{\sigma}} e^{-\beta E_{XOR}(\underline{\sigma})} \times \log \left[ \sum_{\underline{\sigma}'} e^{-\beta E_{XOR}(\underline{\sigma}')} \delta(q - q_s(\underline{\sigma}, \underline{\sigma}')) \right], \quad (18)$$

where  $q_s(\underline{\sigma}, \underline{\sigma}') = \frac{1}{N} \sum_{i=1}^N \sigma_i \sigma'_i$  is the spin overlap function that measures the similarity of two configurations,  $\sigma$  and  $\sigma'$ . The potential  $V_s^{XOR}(q)$  represents the free energy of a system of spins  $\underline{\sigma}'$  constrained to be at fixed spin

overlap  $q$  from the spin system defined by  $\underline{\sigma}$ . For mean-field models described by the RFOT theory, the potential  $V_s^{XOR}(q)$  has the following properties, see Fig. 2:

- It is a convex function with a single minimum at  $q = 0$  for  $T > T_d$
- For  $T_d > T > T_k$  it displays a secondary minimum at  $\bar{q}_s^{XOR} > 0$ , such that  $V_s^{XOR}(\bar{q}_s^{XOR}) > V_s^{XOR}(0)$ . The difference between the two minima can be interpreted as the free energy cost to force the two system to have a large overlap.
- For  $T < T_k$  the secondary minimum becomes thermodynamically stable,  $V_s^{XOR}(\bar{q}_s^{XOR}) < V_s^{XOR}(0)$ , signaling the onset of replica symmetry breaking.

For the model under consideration, i.e. XORSAT with  $p = c$ , it turns out that  $T_k = 0$ , as we already discussed; thus the third regime is not relevant for the present discussion. Importantly enough, the value of the overlap  $\bar{q}_s^{XOR}$  where the secondary minimum is attained when  $T < T_d$  represents the plateau at which the dynamical correlations of the spins converge in the long time limit. Thus, the non-trivial dynamical behavior can be understood in terms of thermodynamical quantities, which is the key feature of RFOT theory [28–30].

## 2. Plaquette overlap

Let us now turn to the discussion of a second observable for the same model. Instead of focusing on spin-spin correlations we study here the dynamics of a “plaquette”

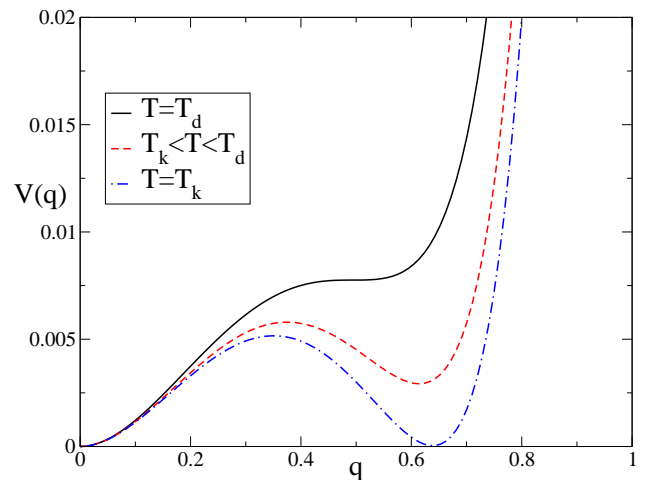


FIG. 2: A schematic representation of the Parisi-Franz potential [28, 29] for a model in the RFOT class.

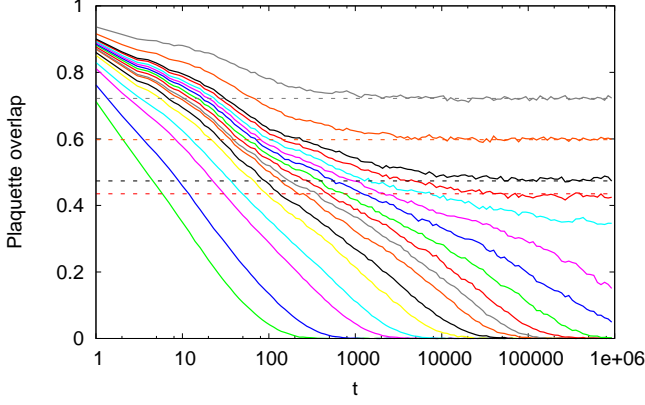


FIG. 3: Local plaquette-plaquette correlation functions  $\phi_p^{XOR}(t)$  (solid lines), and the corresponding “intra-state” thermodynamical overlap (dashed lines) computed with the cavity method [25]. Different curves represent the same temperatures as in Fig. 1.

variable made of three interacting spins, i.e. we consider:

$$C_p^{XOR}(t) = \frac{1}{N} \sum_{a=1}^N \langle \sigma_{i_a^1}(t) \sigma_{i_a^2}(t) \sigma_{i_a^3}(t) \sigma_{i_a^1}(0) \sigma_{i_a^2}(0) \sigma_{i_a^3}(0) \rangle, \\ \langle \sigma_{i_a^1}(t) \sigma_{i_a^2}(t) \sigma_{i_a^3}(t) \sigma_{i_a^1}(0) \sigma_{i_a^2}(0) \sigma_{i_a^3}(0) \rangle = \\ = \langle GS_{XOR} | \hat{\sigma}_{i_a^1}^z \hat{\sigma}_{i_a^2}^z \hat{\sigma}_{i_a^3}^z e^{-\hat{H}_{XOR} t} \hat{\sigma}_{i_a^1}^z \hat{\sigma}_{i_a^2}^z \hat{\sigma}_{i_a^3}^z | GS_{XOR} \rangle. \quad (19)$$

Because the average of a plaquette variable is  $\langle \sigma_{i_a^1} \sigma_{i_a^2} \sigma_{i_a^3} \rangle = \tanh(\beta)$ , it is better to consider the connected and normalized function

$$\phi_p^{XOR}(t) = \frac{C_p^{XOR}(t) - \tanh^2(\beta)}{C_p^{XOR}(0) - \tanh^2(\beta)}. \quad (20)$$

This correlation is shown in Fig. 3 for different temperatures; it displays a behavior slightly different from the one shown in Fig. 1, but still below the dynamical transition  $T < T_d$  it does never relax and it reaches a plateau whose value can be computed through the cavity method [25], similarly to what done in the previous case. This plateau value represents the overlap of plaquettes of spins between configurations belonging to the same state. This behavior is not surprising in light of the understanding of the dynamical transition as the decomposition of the measure in many states, within the RFOT theory picture. With such local dynamics, the system is not able to escape from the initial state and the cavity method allows us to access precisely this intra-state properties.

Because the dynamical behavior of the spin and plaquette overlap is qualitatively equivalent, one would like to understand whether, from a thermodynamic point of view, the plaquette overlap can be used as an order parameter to constrain the system in one metastable state,

through the Franz-Parisi potential. For the plaquette overlap, the Franz-Parisi potential  $V_p^{XOR}(q)$  is defined as in (18), with  $q_s(\underline{q}, \underline{q}')$  replaced by

$$q_p(\underline{q}, \underline{q}') = \frac{1}{N} \sum_{a=1}^N \sigma_{i_a^1} \sigma_{i_a^2} \sigma_{i_a^3} \sigma'_{i_a^1} \sigma'_{i_a^2} \sigma'_{i_a^3}. \quad (21)$$

An explicit calculation of this quantity is possible, by exploiting the mapping towards the non-interacting model discussed in Section II C, thanks to the crucial fact that the plaquette overlap in XORSAT becomes the spin overlap in the KCM. Therefore  $V_p^{XOR}(q) = V_s^{KCM}(q)$ . The computation is done in Appendix A and it shows explicitly that the function  $V_p^{XOR}(q)$  is a convex function at all temperatures, with a single minimum at  $q = \tanh^2(\beta)$  corresponding to the thermodynamic value of  $q_p$  in the paramagnetic state. We conclude that  $V_p^{XOR}(q)$  does not signal the presence of a metastable state and not even provide an interpretation for the change in the dynamical behavior of the plaquette correlation in Eq. (19) when  $T < T_d$ .

This example shows that in order to find a satisfactory thermodynamic description that accounts for the dynamical behavior it is important to select the correct order parameter which in the case of the XORSAT model is the spin overlap. Even if, a priori, imposing a given spin overlap between two configurations might not seem very different from fixing the plaquette overlap, this leads to a qualitative different behavior of the corresponding Franz-Parisi potential. The reason is readily found by thinking to the structure of the energy landscape of XORSAT, and in particular, as in the case under study, when there is a single ground state. Up to an overall constant the ground state has energy zero and above it there are  $N$  excited configurations that violate exactly one interaction (i.e. one plaquette has value  $-1$ ). These configurations are at extensive Hamming distance among each other and with the ground state, therefore their spin overlap with the ground state is very close to zero. However, in terms of plaquettes, they are at distance 1 from the ground state and at distance 2 among each other, therefore their plaquette overlap with the ground state and among each other is very large. This implies that the requirement to fix a high value of the plaquette overlap in the Franz-Parisi potential does not select a state and on the contrary allows to sum over configurations belonging to different states. For this reason the overlap  $q_p(\underline{q}, \underline{q}')$  does not work as an order parameter that signals the emergence of many metastable states that trap the dynamics. This was also recognized in the context of finite-dimensional plaquette models [12].

The situation is somehow analogous to what happens if one considers a ferromagnetic Ising model and adds an infinitesimal field which couples to  $s_i s_j$ . Clearly, this field does not break the symmetry and is therefore unable to detect the phase transition by constraining the system in one of the two symmetry-breaking pure states. However, the dynamical correlations of  $s_i s_j$  would clearly show a

plateau in the low temperature phase if the system is constrained into one of its pure states.

### B. Kinetically constrained model

We turn now to the discussion of the kinetically constrained model, introduced in Section II B. Due to the mapping onto XORSAT, non-trivial properties in the dynamics of the KCM are expected to appear in spite of the simplicity of the energy function Eq. (8).

The equilibrium local spin-spin correlations for this model read:

$$\begin{aligned}
 C_s^{KCM}(t) &= \frac{1}{N} \sum_{a=1}^N \langle s_a(t) s_a(0) \rangle, \\
 \langle s_a(t) s_a(0) \rangle &= \langle GS_{KCM} | \hat{s}_a^x e^{-\hat{H}_{KCM} t} \hat{s}_a^x | GS_{KCM} \rangle, \\
 \phi_s^{KCM}(t) &= \frac{C_s^{KCM}(t) - \tanh^2(\beta)}{C_s^{KCM}(0) - \tanh^2(\beta)},
 \end{aligned} \tag{22}$$

where  $\langle \underline{s} | GS_{KCM} \rangle$  is the ground state of  $\hat{H}_{KCM}$  defined in Eq. (14). Thanks to the duality transformation the ground state of  $\hat{H}_{KCM}$  is mapped into that of  $\hat{H}_{XOR}$  and the time evolution operator as well. Moreover the observable  $\hat{s}_a^x$  is mapped in  $\hat{\sigma}_{i_1}^z \hat{\sigma}_{i_2}^z \hat{\sigma}_{i_3}^z$  in the XORSAT problem. Thus, the spin correlation in the KCM, Eq. (22) is equivalent to the plaquette correlation in the XORSAT problem, Eq. (19):  $C_s^{KCM}(t) = C_p^{XOR}(t)$ . This equivalence is confirmed numerically in Fig. 4 at two different temperatures,  $T = 0.45$  and  $T = 0.6$ . The correlation functions of the KCM have been obtained using Metropolis dynamics in discrete time for a system with  $N = 60000$  spins. In the case of the KCM, the problem of generating the initial equilibrium configuration is of course completely trivial.

Thus, the dynamics of the spins in the KCM is the same shown in Fig. 3 for the plaquettes in XORSAT. It reveals non-trivial features and in particular a dynamical transition at the same critical temperature  $T_d = 0.5098(2)$ . Such a transition cannot be understood in terms of a simple thermodynamical reasoning on the basis of the Franz-Parisi potential for the spin overlap of the KCM. Indeed, as we already discussed,  $V_s^{KCM}(q) = V_p^{XOR}(q)$  and both are trivial functions of  $q$ , see Appendix A. Indeed, due to the non-interacting nature of the spins of the KCM, from the static point of view the overlap  $q_s(\underline{s}, \underline{s}') = \frac{1}{N} \sum_{a=1}^N s_a s'_a$  is not expected to select any state and it cannot work as a static order parameter. This phenomenology is very reminiscent of other kinetically constrained models. At first sight thus, in the KCM one is not able to understand the dynamic transition as the consequence of a complex thermodynamic behavior.

However, the mapping in Eq. (12) teaches us that *there is a static order parameter that does this work*, i.e. what corresponds to the overlap between spin configurations

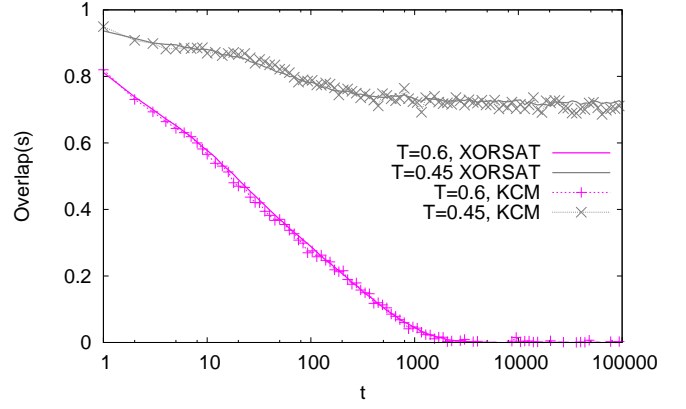


FIG. 4: Comparison between the (connected and normalized) spin-spin correlation function of the KCM,  $\phi_s^{KCM}(t)$  (points), with the plaquette-plaquette correlation  $\phi_p^{XOR}(t)$  of XORSAT (solid lines). The grey and the pink curves refer to different temperatures, respectively  $T = 0.6$  and  $T = 0.45$ .

of the XORSAT model. Expressed in terms of the spins  $s_a$ , it becomes:

$$q_{nl}(\underline{s}, \underline{s}') = \frac{1}{N} \sum_{i=1}^N \prod_{a=1}^N (s_a s'_a)^{y_i^a}. \tag{23}$$

However this quantity is a highly non local observable and then it cannot be easily and naturally recognized, especially if the mapping is not explicitly given (note that in particular the  $y_i^a$  depend on the particular instance of random XORSAT under consideration).

## IV. DISCUSSION

### A. Summary of the results for the Bethe lattice model

In this work we have shown that it is possible to establish a mapping between the equilibrium dynamics of a glassy mean-field model described by the Random First-Order Transition theory, i.e. the XORSAT model, and a Kinetically Constrained Model.

We have discussed the role of the overlap function that is chosen in the framework of RFOT theory to work as order parameter and signal the presence of metastable states. Overlap functions that couple different observables define different distances between configurations and thus might fail to select a particular state. This is the case of the plaquette overlap in the XORSAT model. This observable is directly mapped into the spin overlap in the KCM. Therefore, the thermodynamic properties, and in particular the Franz-Parisi potential, are trivial when the plaquette overlap is considered in XORSAT, or the spin overlap is considered in the KCM.

However the dynamics of all local correlations shows no relaxation below a critical temperature  $T_d$ . Within the XORSAT model, this is signaled by a static computation of the Franz-Parisi potential for the spin overlap, as usual in RFOT theory. For the KCM, thanks to the mapping back to the XORSAT model, one can recast the non-trivial dynamical behavior within a thermodynamical description whose basic degrees of freedom however are hidden non-local functions of the spins.

Thus, this example illustrates that –at least in this case– one is able to resolve the apparent dichotomy that exist between static and dynamic properties tracing them back to the physics of RFOT theory. This is not obvious from the point of view of the KCM: if one were to study it without any knowledge on the mapping on the RFOT model, identifying the correct order parameter would be impossible.

### B. Finite dimensional models

As we discussed in the introduction, the mapping we used for the Bethe lattice model was already used extensively in a class of finite dimensional plaquette models [5, 11–13]. However, these plaquette models do not seem to be described by the RFOT theory, because the energy cost due to the presence of an interface between two low energy states does not grow with system size [11–13, 31].

On the other hand, we have shown in section II that the mapping between the  $p$ -spin model and the KCM holds quite generally, under the only requirements that

*i)* each spin is involved in exactly  $p$  interactions, *ii)* each interaction involves exactly  $p$  spin, and *iii)* the ground state is unique. Therefore, it might be possible to devise a finite-dimensional model (possibly in three or higher dimensions) that satisfies these requirements and at the same time is well described by the RFOT scenario. This could be possibly achieved by extensions of the so-called goniheric models [32, 33] or of the three dimensional plaquette model considered in [34]. A numerical computation of the Franz-Parisi potential for the spin overlap in these models, following the strategy of [35, 36], could be instructive.

Such a model could be quite useful to investigate the relation between the RFOT and KCM pictures in a finite dimensional setting. We recall, however, that the mapping implies that  $T_k = 0$ , because the free energy must be analytic at all temperatures, therefore a RFOT scenario with a positive Kauzmann temperature would not fit in this analysis.

Whether a similar mapping can be constructed in more general models remains an open problem [22]. Finally, it would be very interesting to check whether a KCM with a continuous transition [37] could be constructed, for which a similar mapping exists.

### Acknowledgments

We thank L. Berthier, G. Biroli, J. P. Garrahan, D. Gosset, R. L. Jack, C. Laumann, V. Martin-Mayor, M. Sellitto and G. Semerjian for many important discussion on this work.

- 
- [1] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, *Phys. Rev. A* **40**, 1045 (1989).
  - [2] T. Castellani and A. Cavagna, *Journal of Statistical Mechanics: Theory and Experiment* **2005**, P05012 (2005).
  - [3] V. Lubchenko and P. G. Wolynes, *Annual Review of Physical Chemistry* **58**, 235 (2007), [arXiv.org:cond-mat/0607349](https://arxiv.org/cond-mat/0607349).
  - [4] G. Biroli and J. Bouchaud, in *Structural Glasses and Supercooled Liquids: Theory, Experiment and Applications*, edited by P.G. Wolynes and V. Lubchenko (Wiley & Sons, 2012), [arXiv:0912.2542](https://arxiv.org/abs/0912.2542).
  - [5] J. Garrahan, *Journal of Physics: Condensed Matter* **14**, 1571 (2002).
  - [6] F. Ritort and P. Sollich, *Advances in Physics* **52**, 219 (2003).
  - [7] J. Garrahan, P. Sollich, and C. Toninelli, in *Dynamical Heterogeneities and Glasses*, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos (Oxford University Press, 2011), [arXiv:1009.6113](https://arxiv.org/abs/1009.6113).
  - [8] A. Cavagna, *Physics Reports* **476**, 51 (2009).
  - [9] L. Berthier and G. Biroli, *Rev. Mod. Phys.* **83**, 587 (2011).
  - [10] S. Whitelam and J. Garrahan, *The Journal of Physical Chemistry B* **108**, 6611 (2004).
  - [11] R. L. Jack, L. Berthier, and J. P. Garrahan, *Phys. Rev. E* **72**, 016103 (2005).
  - [12] R. L. Jack and J. P. Garrahan, *The Journal of Chemical Physics* **123**, 164508 (pages 13) (2005).
  - [13] C. Cammarota and G. Biroli (2012), [arXiv:1201.2164](https://arxiv.org/abs/1201.2164).
  - [14] D. Gosset, *Phd thesis* (2010).
  - [15] S. Franz, M. Mézard, F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, *EPL (Europhysics Letters)* **55**, 465 (2001).
  - [16] S. Franz, M. Leone, F. Ricci-Tersenghi, and R. Zecchina, *Phys. Rev. Lett.* **87**, 127209 (2001).
  - [17] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, *J. Stat. Phys.* **111**, 505 (2003).
  - [18] S. Cocco, O. Dubois, J. Mandler, and R. Monasson, *Phys. Rev. Lett.* **90**, 047205 (2003).
  - [19] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, *Phys. Rev. E* **63**, 026702 (2001).
  - [20] A. Montanari and G. Semerjian, *J. Stat. Phys.* **124**, 103 (2006).
  - [21] M. Sellitto, G. Biroli, and C. Toninelli, *Europhysics Letters (EPL)* **69**, 496 (2005).
  - [22] J. Kurchan, L. Peliti, and M. Sellitto, *EPL (Europhysics Letters)* **39**, 365 (1997).
  - [23] T. Jörg, F. Krzakala, G. Semerjian, and F. Zamponi, *Phys. Rev. Lett.* **104**, 207206 (2010).

- [24] L. Zdeborová and M. Mézard, Phys. Rev. Lett. **101**, 078702 (2008).  
 [25] Krzakala, F. and Zdeborová, L., EPL **90**, 66002 (2010).  
 [26] F. Krzakala and L. Zdeborová, Phys. Rev. Lett. **102**, 238701 (2009).  
 [27] L. Zdeborová and F. Krzakala, Phys. Rev. B **81**, 224205 (2010).  
 [28] S. Franz and G. Parisi, Journal de Physique I **5**, 1401 (1995).  
 [29] S. Franz and G. Parisi, Phys. Rev. Lett. **79**, 2486 (1997).  
 [30] S. Franz, G. Parisi, F. Ricci-Tersenghi, and T. Rizzo, The European Physical Journal E: Soft Matter and Biological Physics **34**, 1 (2011).  
 [31] R. L. Jack and L. Berthier, Phys. Rev. E **85**, 021120 (2012).  
 [32] A. Lipowski and D. Johnston, Phys. Rev. E **61**, 6375 (2000).  
 [33] P. Dimopoulos, D. Espriu, E. Jané, and A. Prats, Phys. Rev. E **66**, 056112 (2002).  
 [34] F. Krzakala and L. Zdeborova, The Journal of Chemical Physics **134**, 034513 (pages 13) (2011).  
 [35] L. Fernandez, V. Martin-Mayor, and D. Yllanes, Nuclear physics B **807**, 424 (2009).  
 [36] C. Cammarota, A. Cavagna, I. Giardina, G. Gradenigo, T. S. Grigera, G. Parisi, and P. Verrocchio, Phys. Rev. Lett. **105**, 055703 (2010).  
 [37] M. Sellitto, D. De Martino, F. Caccioli, and J. J. Arenzon, Phys. Rev. Lett. **105**, 265704 (2010).

### Appendix A: Franz-Parisi potential for independent spins

Here we report the calculation of the Franz-Parisi potential for the plaquette overlap in XORSAT,  $V_p^{XOR}(q)$ , that is mapped onto the same potential for the spin overlap in the KCM,  $V_s^{KCM}(q)$ . Since in the KCM the spins are independent, the latter is trivially computed. We have:

$$V_p^{XOR}(q) = V_s^{KCM}(q) = -\frac{1}{\beta Z} \sum_{\underline{\sigma}} e^{-\beta E_{KCM}(\underline{\sigma})} \log \left[ \sum_{\underline{\sigma}'} e^{-\beta E_{KCM}(\underline{\sigma}')} \delta(q - q_s(\underline{\sigma}, \underline{\sigma}')) \right] \quad (\text{A1})$$

where the spin overlap is  $q_s(\underline{\sigma}, \underline{\sigma}') = \frac{1}{N} \sum_i \sigma_i \sigma'_i$  for a system of free spins in magnetic field  $E_{KCM}(\underline{\sigma}) = -\sum_i \sigma_i$ . Exploiting the relation  $\log x = \lim_{n \rightarrow 0} \partial_n x^n$  the Franz-Parisi potential becomes:

$$\begin{aligned} V_s^{KCM}(q) &= -\lim_{n \rightarrow 0} \partial_n \sum_{\underline{\tau}} \frac{e^{\beta \sum_i \tau_i}}{\beta (2 \cosh \beta)^N} \sum_{\tau^1, \dots, \tau^n} e^{\beta \sum_{i,a} \tau_i^a} \prod_{a=1}^n \delta \left( Nq - \sum_i \tau_i \tau_i^a \right) \\ &= -\lim_{n \rightarrow 0} \partial_n \frac{1}{\beta (2 \cosh \beta)^N} \int d\lambda_1 \dots d\lambda_n e^{-\beta Nq \sum_{a=1}^n \lambda_a} \sum_{\underline{\tau}, \tau^1, \dots, \tau^n} e^{\beta (\sum_i \tau_i + \sum_{i,a} \tau_i^a) + \beta \sum_{i,a} \lambda_a \tau_i \tau_i^a} \\ &= -\lim_{n \rightarrow 0} \partial_n \frac{1}{\beta (2 \cosh \beta)^N} \int d\lambda_1 \dots d\lambda_n e^{-\beta Nq \sum_{a=1}^n \lambda_a} \left( \sum_{\tau, \tau^1, \dots, \tau^n} e^{\beta (\tau + \sum_a \tau^a) + \beta \sum_a \lambda_a \tau \tau^a} \right)^N \\ &= -\lim_{n \rightarrow 0} \partial_n \frac{1}{\beta (2 \cosh \beta)^N} \int d\lambda e^{-\beta Nqn\lambda + N \log f(\lambda)} \end{aligned} \quad (\text{A2})$$

with

$$\begin{aligned} f(\lambda) &= \sum_{\tau, \tau^1, \dots, \tau^n} e^{\beta (\tau + \sum_a \tau^a) + \beta \lambda \sum_a \tau \tau^a} = \sum_{\tau} e^{\beta \tau} \left[ 2 \cosh \left( \beta (1 + \lambda \tau) \right) \right]^n \\ &= e^{\beta} \left[ 2 \cosh \left( \beta (1 + \lambda) \right) \right]^n + e^{-\beta} \left[ 2 \cosh \left( \beta (1 - \lambda) \right) \right]^n \end{aligned} \quad (\text{A3})$$

Developing for small  $n$

$$f(\lambda) = 2 \cosh(\beta) \left[ 1 + \frac{n}{2 \cosh \beta} (e^{-\beta} \log [2 \cosh(\beta(1-\lambda))] + e^{\beta} \log [2 \cosh(\beta(1+\lambda))]) \right] \quad (\text{A4})$$

Thus

$$\begin{aligned} V_s^{KCM}(q) &= -\frac{1}{\beta} \lim_{n \rightarrow 0} \partial_n \int d\lambda e^{-\beta Nqn\lambda + \frac{Nn}{2 \cosh \beta} (e^{-\beta} \log [2 \cosh(\beta(1-\lambda))] + e^{\beta} \log [2 \cosh(\beta(1+\lambda))])} \\ &\simeq \frac{N}{\beta} \min_{\lambda} \left[ \beta q \lambda - \frac{1}{2 \cosh \beta} (e^{-\beta} \log [2 \cosh(\beta(1-\lambda))] + e^{\beta} \log [2 \cosh(\beta(1+\lambda))]) \right] \end{aligned} \quad (\text{A5})$$

and

$$q(\lambda) = \frac{1}{2 \cosh \beta} \left( e^\beta \tanh(\beta(1 + \lambda)) - e^{-\beta} \tanh(\beta(1 - \lambda)) \right) \quad (\text{A6})$$

The function (A5) is a convex function at all temperatures, with a single minimum at  $q^* = q(\lambda = 0) = (\tanh \beta)^2$ , such that  $V_s^{KCM}(q^*) = F(T) = -\frac{N}{\beta} \log 2 \cosh(\beta)$ .

---